

Current protocols example dataset

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 4.4.3(git:) [March-21-2019]

Using Intel AVX (Advanced Vector Extensions)

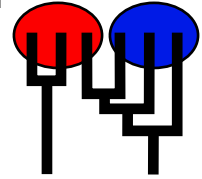
Compiled for PARALLEL computer architectures

One master and 8 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Sat Jun 1 22:21:29 2019

Program finished at Sat Jun 1 22:49:35 2019 [Runtime:0000:00:28:06]



Options

Inheritance scalers in use for Thetas:

All loci use an inheritance scaler of 1.0

[The locus with a scaler of 1.0 used as reference]

Random number seed: (with internal timer) 2162460700

Start parameters:

Theta values were generated Using a percent value of the prior

M values were generated Using a percent value of the prior

Connection matrix:

m = average (average over a group of Thetas or M,

s = symmetric migration M, S = symmetric 4Nm,

0 = zero, and not estimated,

* = migration free to vary, Thetas are on diagonal

d = row population split off column population, D = split and then migration

Population	1	2	3
1 Arbon_1	*	*	*
2 Berg_2	*	*	*
3 Chur_3	*	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	M _{2→1}	<displayed>
5	M _{3→1}	<displayed>
6	M _{1→2}	<displayed>
7	M _{3→2}	<displayed>
8	M _{1→3}	<displayed>
9	M _{2→3}	<displayed>

Mutation rate among loci:

Mutation rate is constant for all loci

Analysis strategy:

Bayesian inference

-Population size estimation:

Exponential Distribution

-Geneflow estimation:

Exponential Distribution

Proposal distributions for parameter

Parameter	Proposal
Theta	Metropolis sampling
M	Metropolis sampling
Divergence	Metropolis sampling
Divergence Spread	Metropolis sampling
Genealogy	Metropolis-Hastings

Prior distribution for parameter

Parameter		Prior	Minimum	Mean	Maximum	Delta	Bins	UpdateFreq
1	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
2	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
3	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
4	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
5	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
6	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
7	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
8	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
9	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.05556

[* * means priors were set globally]

Markov chain settings:

Long chain

Number of chains	1
Recorded steps [a]	5000
Increment (record every x step [b])	100
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	500000
Number of discard trees per chain (burn-in)	5000

Multiple Markov chains:

Static heating scheme

1000000.00 4 chains with temperatures
3.00 1.50 1.00
Swapping interval is 1

Print options:

Data file: infile
Haplotyping is turned on: NO
Output file: outfile
Posterior distribution raw histogram file: bayesfile
Raw data from the MCMC run: bayesallfile.gz
Print data: No
Print genealogies [only some for some data type]: None

Data summary

Data file: infile
 Datatype: Haplotype data
 Number of loci: 10

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Jukes-Cantor	[Basefreq: =0.25]
2	1	Jukes-Cantor	[Basefreq: =0.25]
3	1	Jukes-Cantor	[Basefreq: =0.25]
4	1	Jukes-Cantor	[Basefreq: =0.25]
5	1	Jukes-Cantor	[Basefreq: =0.25]
6	1	Jukes-Cantor	[Basefreq: =0.25]
7	1	Jukes-Cantor	[Basefreq: =0.25]
8	1	Jukes-Cantor	[Basefreq: =0.25]
9	1	Jukes-Cantor	[Basefreq: =0.25]
10	1	Jukes-Cantor	[Basefreq: =0.25]

Sites per locus

Locus	Sites
1	1000
2	1000
3	1000
4	1000
5	1000
6	1000
7	1000
8	1000
9	1000
10	1000

Site rate variation and probabilities:

Locus	Sublocus	Region type	Rate of change	Probability	Patch size
1	1	1	1.000	1.000	1.000
2	1	1	1.000	1.000	1.000
3	1	1	1.000	1.000	1.000
4	1	1	1.000	1.000	1.000
5	1	1	1.000	1.000	1.000
6	1	1	1.000	1.000	1.000

7	1	1	1.000	1.000	1.000		
8	1	1	1.000	1.000	1.000		
9	1	1	1.000	1.000	1.000		
10	1	1	1.000	1.000	1.000		
Population				Locus		Gene copies data	(missing)
1 Arbon_1				1		10	
				2		10	
				3		10	
				4		10	
				5		10	
				6		10	
				7		10	
				8		10	
				9		10	
				10		10	
2 Berg_2				1		10	
				2		10	
				3		10	
				4		10	
				5		10	
				6		10	
				7		10	
				8		10	
				9		10	
				10		10	
3 Chur_3				1		10	
				2		10	
				3		10	
				4		10	
				5		10	
				6		10	
				7		10	
				8		10	
				9		10	
				10		10	
Total of all populations				1		30	(0)
				2		30	(0)
				3		30	(0)
				4		30	(0)
				5		30	(0)
				6		30	(0)
				7		30	(0)
				8		30	(0)
				9		30	(0)

10

30

(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00180	0.00427	0.00503	0.00620	0.01907	0.01137	0.02142
1	Θ_2	0.00093	0.00193	0.00310	0.00353	0.00813	0.00370	0.00476
1	Θ_3	0.00127	0.00413	0.00423	0.00520	0.01513	0.00737	0.00903
1	$M_{2 \rightarrow 1}$	253.333	696.667	705.000	706.667	1066.667	728.333	750.594
1	$M_{3 \rightarrow 1}$	116.667	510.000	528.333	580.000	1080.000	598.333	648.945
1	$M_{1 \rightarrow 2}$	76.667	483.333	498.333	520.000	860.000	498.333	548.762
1	$M_{3 \rightarrow 2}$	43.333	186.667	218.333	220.000	580.000	285.000	307.931
1	$M_{1 \rightarrow 3}$	303.333	503.333	525.000	526.667	1056.667	641.667	676.760
1	$M_{2 \rightarrow 3}$	0.000	0.000	28.333	70.000	363.333	148.333	259.915
2	Θ_1	0.00160	0.00260	0.00383	0.00500	0.00840	0.00443	0.00488
2	Θ_2	0.00027	0.00120	0.00197	0.00360	0.01093	0.00370	0.00957
2	Θ_3	0.00073	0.00213	0.00277	0.00380	0.00793	0.00370	0.00423
2	$M_{2 \rightarrow 1}$	0.000	16.667	65.000	103.333	346.667	105.000	137.886
2	$M_{3 \rightarrow 1}$	0.000	10.000	48.333	103.333	343.333	95.000	130.658
2	$M_{1 \rightarrow 2}$	0.000	0.000	45.000	83.333	836.667	628.333	838.651
2	$M_{3 \rightarrow 2}$	3.333	126.667	248.333	270.000	766.667	358.333	455.960
2	$M_{1 \rightarrow 3}$	60.000	160.000	208.333	286.667	616.667	328.333	385.710
2	$M_{2 \rightarrow 3}$	0.000	0.000	1.667	33.333	206.667	55.000	76.772
3	Θ_1	0.00120	0.00433	0.00517	0.00533	0.00967	0.00550	0.00574
3	Θ_2	0.00127	0.00460	0.00650	0.00767	0.02127	0.02630	0.03493
3	Θ_3	0.00293	0.00647	0.00717	0.00767	0.01560	0.00883	0.00970
3	$M_{2 \rightarrow 1}$	0.000	0.000	5.000	50.000	263.333	61.667	128.024
3	$M_{3 \rightarrow 1}$	0.000	33.333	38.333	60.000	256.667	81.667	122.667
3	$M_{1 \rightarrow 2}$	23.333	80.000	108.333	173.333	173.333	2565.000	2507.317
3	$M_{3 \rightarrow 2}$	0.000	0.000	1.667	156.667	450.000	211.667	398.983
3	$M_{1 \rightarrow 3}$	0.000	0.000	1.667	30.000	140.000	31.667	47.296
3	$M_{2 \rightarrow 3}$	0.000	0.000	1.667	26.667	123.333	31.667	45.896
4	Θ_1	0.00080	0.00207	0.00237	0.00293	0.00840	0.00390	0.00449
4	Θ_2	0.00160	0.00293	0.00350	0.00480	0.00887	0.00457	0.00496
4	Θ_3	0.00233	0.00473	0.00497	0.00827	0.01560	0.00810	0.00913
4	$M_{2 \rightarrow 1}$	0.000	30.000	101.667	140.000	450.000	168.333	205.334
4	$M_{3 \rightarrow 1}$	166.667	296.667	305.000	306.667	893.333	558.333	606.140
4	$M_{1 \rightarrow 2}$	3.333	23.333	35.000	66.667	170.000	71.667	81.813

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$M_{3 \rightarrow 2}$	0.000	16.667	58.333	70.000	166.667	65.000	73.488
4	$M_{1 \rightarrow 3}$	46.667	136.667	191.667	223.333	460.000	225.000	243.440
4	$M_{2 \rightarrow 3}$	26.667	83.333	111.667	173.333	323.333	155.000	166.951
5	Θ_1	0.00080	0.00120	0.00137	0.00287	0.02107	0.02137	0.03304
5	Θ_2	0.00133	0.00180	0.00263	0.00267	0.00660	0.00350	0.00378
5	Θ_3	0.00573	0.00913	0.00950	0.00980	0.02393	0.01410	0.01555
5	$M_{2 \rightarrow 1}$	893.333	893.333	905.000	913.333	1060.000	1605.000	1805.062
5	$M_{3 \rightarrow 1}$	6.667	183.333	255.000	263.333	880.000	425.000	525.674
5	$M_{1 \rightarrow 2}$	0.000	0.000	8.333	36.667	250.000	65.000	97.579
5	$M_{3 \rightarrow 2}$	0.000	16.667	51.667	110.000	263.333	95.000	110.232
5	$M_{1 \rightarrow 3}$	0.000	3.333	28.333	70.000	210.000	65.000	83.033
5	$M_{2 \rightarrow 3}$	0.000	0.000	1.667	23.333	103.333	25.000	35.266
6	Θ_1	0.00073	0.00213	0.00277	0.00360	0.01660	0.00530	0.01206
6	Θ_2	0.00093	0.00173	0.00303	0.00333	0.00707	0.00323	0.00373
6	Θ_3	0.00467	0.01273	0.01283	0.01347	0.02993	0.01603	0.02159
6	$M_{2 \rightarrow 1}$	1546.667	1573.333	1585.000	1590.000	1696.667	1725.000	1888.815
6	$M_{3 \rightarrow 1}$	0.000	110.000	135.000	163.333	773.333	338.333	447.474
6	$M_{1 \rightarrow 2}$	0.000	206.667	235.000	270.000	820.000	455.000	654.861
6	$M_{3 \rightarrow 2}$	0.000	0.000	5.000	50.000	350.000	128.333	172.826
6	$M_{1 \rightarrow 3}$	0.000	0.000	11.667	150.000	760.000	231.667	348.537
6	$M_{2 \rightarrow 3}$	0.000	0.000	5.000	6.667	550.000	201.667	265.521
7	Θ_1	0.00200	0.00320	0.00463	0.00480	0.01093	0.00537	0.00608
7	Θ_2	0.00327	0.00427	0.00463	0.00487	0.02593	0.02537	0.03429
7	Θ_3	0.00373	0.01033	0.01097	0.01107	0.02260	0.01590	0.02462
7	$M_{2 \rightarrow 1}$	76.667	233.333	245.000	310.000	640.000	331.667	373.928
7	$M_{3 \rightarrow 1}$	0.000	33.333	61.667	73.333	253.333	91.667	111.918
7	$M_{1 \rightarrow 2}$	656.667	903.333	908.333	916.667	1646.667	1155.000	1210.422
7	$M_{3 \rightarrow 2}$	0.000	36.667	41.667	46.667	400.000	148.333	194.748
7	$M_{1 \rightarrow 3}$	0.000	23.333	108.333	176.667	760.000	295.000	522.954
7	$M_{2 \rightarrow 3}$	0.000	40.000	51.667	173.333	530.000	205.000	312.852
8	Θ_1	0.00093	0.00353	0.00363	0.00387	0.00807	0.00417	0.00496
8	Θ_2	0.00100	0.00267	0.00323	0.00433	0.01093	0.00490	0.00616
8	Θ_3	0.00193	0.00520	0.00543	0.00600	0.01933	0.00930	0.01070
8	$M_{2 \rightarrow 1}$	0.000	0.000	8.333	60.000	483.333	168.333	353.573
8	$M_{3 \rightarrow 1}$	0.000	73.333	88.333	140.000	476.667	165.000	252.311
8	$M_{1 \rightarrow 2}$	3.333	186.667	238.333	310.000	1463.333	815.000	1061.030
8	$M_{3 \rightarrow 2}$	0.000	26.667	38.333	63.333	546.667	228.333	299.790

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$M_{1 \rightarrow 3}$	0.000	0.000	1.667	33.333	133.333	35.000	50.664
8	$M_{2 \rightarrow 3}$	0.000	0.000	11.667	26.667	236.667	58.333	89.832
9	Θ_1	0.00147	0.00227	0.00297	0.00400	0.01867	0.01070	0.02299
9	Θ_2	0.00180	0.00373	0.00403	0.00507	0.00880	0.00490	0.00524
9	Θ_3	0.00687	0.00813	0.00863	0.00867	0.02833	0.02690	0.03646
9	$M_{2 \rightarrow 1}$	970.000	1180.000	1191.667	1193.333	2000.000	1468.333	1502.752
9	$M_{3 \rightarrow 1}$	0.000	50.000	81.667	113.333	410.000	151.667	180.210
9	$M_{1 \rightarrow 2}$	133.333	293.333	321.667	343.333	563.333	345.000	358.498
9	$M_{3 \rightarrow 2}$	213.333	336.667	378.333	426.667	686.667	421.667	435.687
9	$M_{1 \rightarrow 3}$	136.667	386.667	391.667	396.667	626.667	378.333	410.923
9	$M_{2 \rightarrow 3}$	210.000	440.000	445.000	510.000	810.000	488.333	507.995
10	Θ_1	0.00113	0.00200	0.00303	0.00387	0.00713	0.00357	0.00398
10	Θ_2	0.00173	0.00360	0.00397	0.00493	0.01007	0.00577	0.00701
10	Θ_3	0.00353	0.00813	0.00843	0.01007	0.02367	0.01397	0.01815
10	$M_{2 \rightarrow 1}$	290.000	520.000	595.000	596.667	1023.333	631.667	654.985
10	$M_{3 \rightarrow 1}$	213.333	620.000	628.333	640.000	1003.333	601.667	630.958
10	$M_{1 \rightarrow 2}$	250.000	480.000	561.667	603.333	1043.333	641.667	681.913
10	$M_{3 \rightarrow 2}$	356.667	693.333	701.667	733.333	1103.333	748.333	767.446
10	$M_{1 \rightarrow 3}$	143.333	300.000	371.667	400.000	696.667	408.333	428.371
10	$M_{2 \rightarrow 3}$	206.667	370.000	388.333	496.667	730.000	451.667	471.048
All	Θ_1	0.00260	0.00340	0.00397	0.00447	0.00553	0.00410	0.00405
All	Θ_2	0.00260	0.00340	0.00390	0.00440	0.00533	0.00403	0.00397
All	Θ_3	0.00547	0.00687	0.00783	0.00873	0.01080	0.00803	0.00810
All	$M_{2 \rightarrow 1}$	170.000	256.667	291.667	323.333	330.000	271.667	268.719
All	$M_{3 \rightarrow 1}$	80.000	120.000	148.333	180.000	256.667	161.667	164.471
All	$M_{1 \rightarrow 2}$	123.333	180.000	218.333	263.333	366.667	235.000	239.527
All	$M_{3 \rightarrow 2}$	63.333	110.000	141.667	170.000	236.667	148.333	149.817
All	$M_{1 \rightarrow 3}$	70.000	110.000	135.000	166.667	230.000	148.333	148.212
All	$M_{2 \rightarrow 3}$	20.000	60.000	81.667	103.333	143.333	85.000	84.482

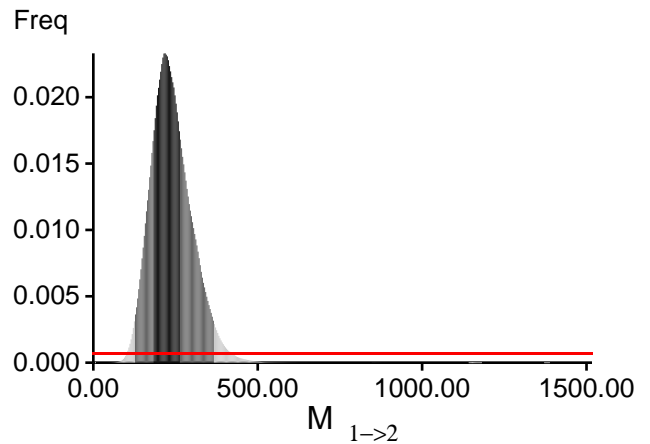
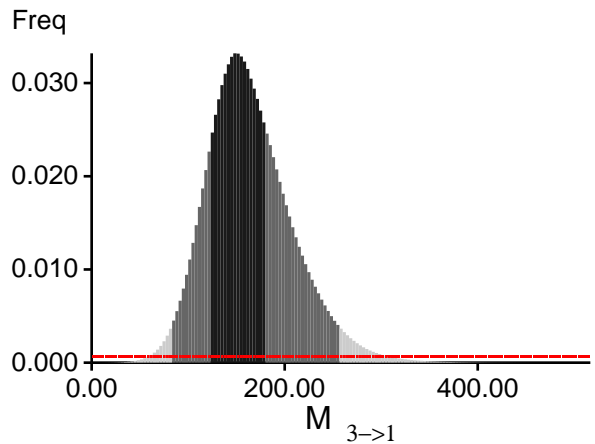
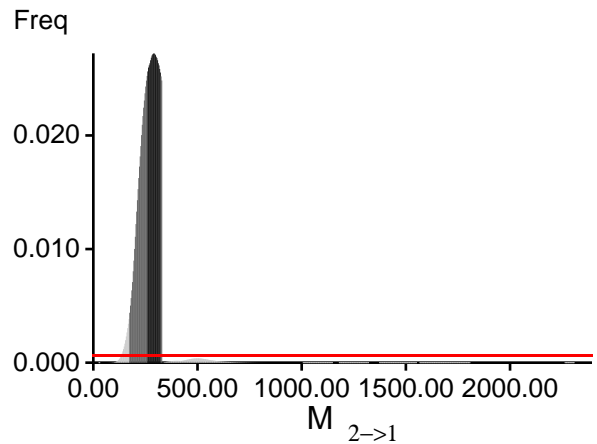
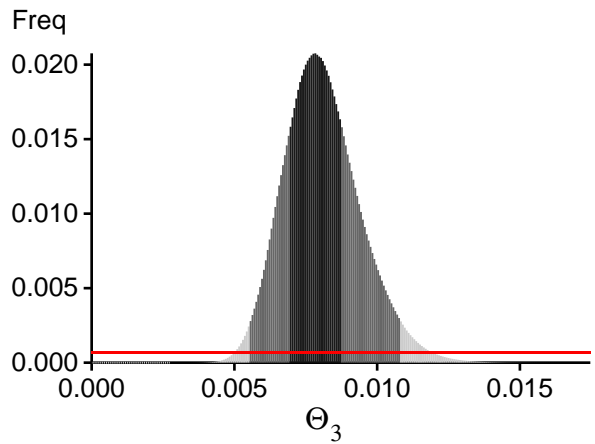
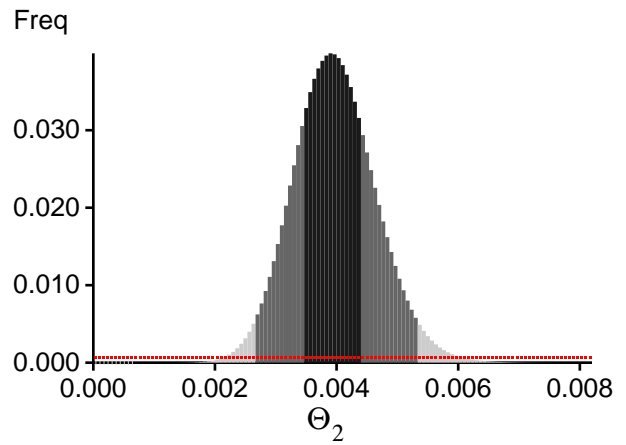
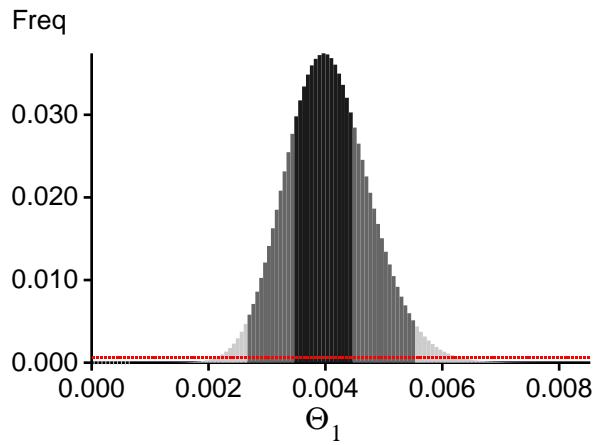
Citation suggestions:

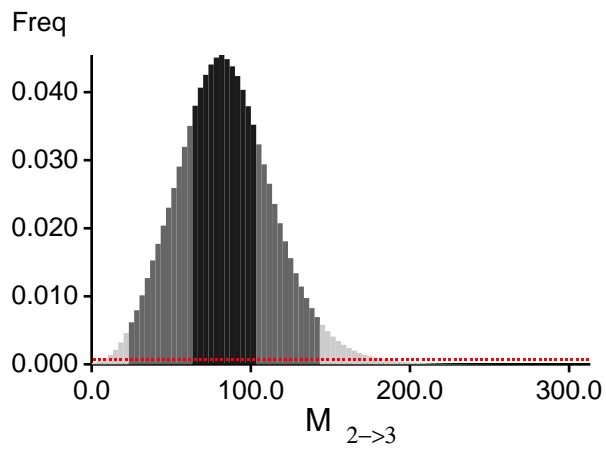
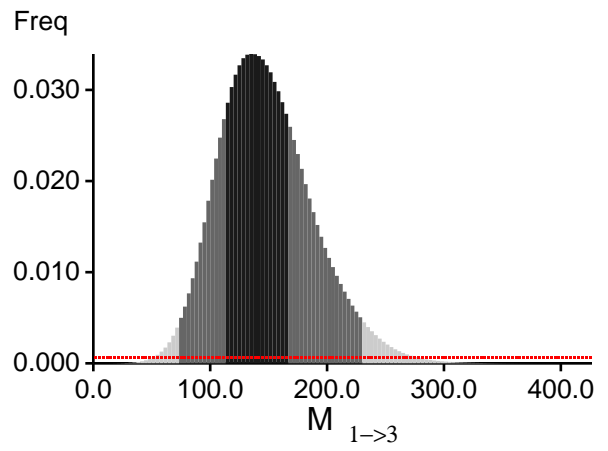
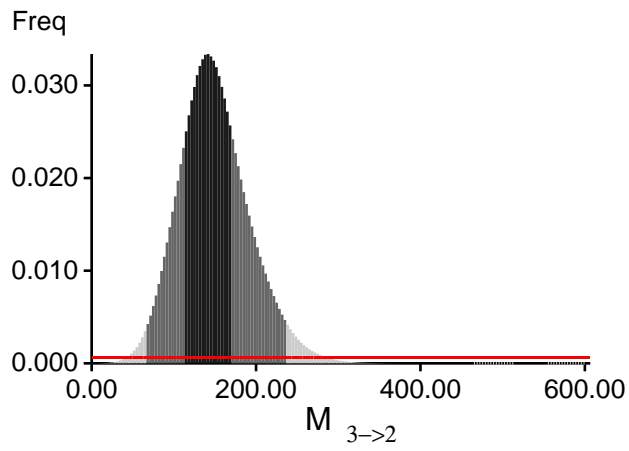
Berli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345

Berli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use?

In *Population Genetics for Animal Conservation*, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of *Conservation Biology*, Cambridge University Press, Cambridge UK, pp. 42-79.

Bayesian Analysis: Posterior distribution over all loci





Log-Probability of the data given the model (marginal likelihood)

Use this value for Bayes factor calculations:

BF = Exp[ln(Prob(D | thisModel) - ln(Prob(D | otherModel)

or as LBF = 2 (ln(Prob(D | thisModel) - ln(Prob(D | otherModel))

shows the support for thisModel]

Locus	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-3259.68	-2481.78	-2316.92
2	-2617.92	-2101.31	-1982.49
3	-2984.41	-2296.79	-2145.04
4	-3402.35	-2451.22	-2252.37
5	-3025.62	-2372.29	-2198.44
6	-3054.68	-2545.68	-2365.67
7	-3063.32	-2403.66	-2217.02
8	-3295.88	-2436.02	-2282.95
9	-3315.35	-2494.27	-2289.66
10	-2983.81	-2316.44	-2159.82
All	-30890.53	-23786.97	-22097.91

(1a, 1b and 2) are approximations to the marginal likelihood, make sure that the program run long enough!

(1a, 1b) and (2) should give similar results, in principle.

But (2) is overestimating the likelihood, it is presented for historical reasons and should not be used

(1a, 1b) needs heating with chains that span a temperature range of 1.0 to at least 100,000.

(1b) is using a Bezier-curve to get better approximations for runs with low number of heated chains

[Scaling factor = 112.486621]

Citation suggestions:

Beerli P. and M. Palczewski, 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations, *Genetics*, 185: 313-326.

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	117341/276666	0.42413
Θ_2	132091/278732	0.47390
Θ_3	121683/277649	0.43826
M _{2→1}	118380/278347	0.42530
M _{3→1}	108032/277931	0.38870
M _{1→2}	143597/278308	0.51596
M _{3→2}	142457/277244	0.51383
M _{1→3}	109581/277469	0.39493
M _{2→3}	137639/277534	0.49594
Genealogies	330262/2500120	0.13210

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.69058	12317.35
Θ_2	0.72734	11513.68
Θ_3	0.61565	16888.79
$M_{2 \rightarrow 1}$	0.72276	11563.76
$M_{3 \rightarrow 1}$	0.62149	16535.95
$M_{1 \rightarrow 2}$	0.74998	11147.20
$M_{3 \rightarrow 2}$	0.61628	18390.41
$M_{1 \rightarrow 3}$	0.68665	13847.61
$M_{2 \rightarrow 3}$	0.70894	13697.48
Genealogies	0.83778	6858.39

Potential Problems

This section reports potential problems with your run, but such reporting is often not very accurate. With many parameters in a multilocus analysis, it is very common that some parameters for some loci will not be very informative, triggering suggestions (for example to increase the prior range) that are not sensible. This suggestion tool will improve with time, therefore do not blindly follow its suggestions. If some parameters are flagged, inspect the tables carefully and judge whether an action is required. For example, if you run a Bayesian inference with sequence data, for macroscopic species there is rarely the need to increase the prior for Theta beyond 0.1; but if you use microsatellites it is rather common that your prior distribution for Theta should have a range from 0.0 to 100 or more. With many populations (>3) it is also very common that some migration routes are estimated poorly because the data contains little or no information for that route. Increasing the range will not help in such situations, reducing number of parameters may help in such situations.

No warning was recorded during the run